

# PENGELOMPOKAN PROVINSI DI INDONESIA BERDASARKAN KRIMINALITAS MENGGUNAKAN METODE WARD DAN K-MEDOIDS

Anggi Dwi Lestari<sup>1</sup>, Tiani Wahyu Utami<sup>2</sup>, M. Al haris<sup>3</sup>

<sup>123</sup>Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang

Alamat e-mail : [anggidwilestari.97@gmail.com](mailto:anggidwilestari.97@gmail.com)

## ABSTRAK

Kejahatan kriminalitas di Indonesia masih menjadi problema yang sulit untuk di tangani karena jumlahnya yang fluktuatif. Badan Pusat Statistik menunjukkan bahwa jumlah kasus kriminalitas di Indonesia selalu meningkat dari tahun ke tahun. Penelitian ini bertujuan untuk mengetahui perbandingan pengklasteran provinsi di Indonesia berdasarkan kriminalitas dengan metode *ward* dan metode *k-medoids*. Objek yang digunakan adalah 34 provinsi di Indonesia dengan 8 variabel indikator kriminalitas tahun 2018. Berdasarkan penelitian yang dilakukan dengan metode *Ward* dan *K-medoids*, kedua metode tersebut menghasilkan pengklasteran yang berbeda. Hasil dari perbandingan kedua metode tersebut diperoleh bahwa metode *Ward* merupakan metode terbaik berdasarkan 4 klaster dengan nilai simpangan baku dalam klaster yaitu 337,7872 dan nilai simpangan baku antar klaster yaitu 1222,236 sehingga menghasilkan nilai rasio simpangan baku yaitu sebesar 0,276368. Hasil Pengklasteran dengan metode *Ward* yaitu Klaster 1 daerah cukup aman yaitu Provinsi Aceh, Jawa Tengah, Kalimantan Selatan, Kalimantan Timur, Lampung, Nusa Tenggara Timur, Papua, Riau, Sulawesi Tengah, dan Sulawesi Utara. Klaster 2 daerah aman yaitu Provinsi Bali, Kepulauan Bangka Belitung, Banten, Bengkulu, Gorontalo, Papua Barat, Jambi, Kalimantan Barat, Kalimantan Tengah, Kalimantan Utara, Kepulauan Riau, Maluku Utara, Maluku, Nusa Tenggara Barat, Sulawesi Barat, Sulawesi Tenggara, dan DI Yogyakarta. Klaster 3 daerah rawan yaitu Provinsi DKI Jakarta dan Sumatera Utara. Dan klaster 4 daerah cukup rawan yaitu Provinsi Jawa Barat, Jawa Timur, Sulawesi Selatan, Sumatera Barat, dan Sumatera Selatan.

**Kata Kunci :** *Ward*, *K-Medoids*, Kriminalitas.

## ABSTRACT

*Crime in Indonesia are still a difficult problem to handle because of the fluctuating numbers. The Central Bureau of Statistics shows that the number of criminal cases in Indonesia is always increasing from year to year. This study aims to determine the comparison of provincial clustering in Indonesia based on crime with the ward method and the method k-medoids. The objects used are 34 provinces in Indonesia with 8 crime indicator variables in 2018. Based on research conducted with the Ward and K-medoids method, the two methods produce different clusters. The results of the comparison of the two methods show that the method Ward is the best method based on 4 clusters with a standard deviation value in the cluster of 337.7872 and a standard deviation value between clusters of 1222.236, resulting in a standard deviation ratio of 0.276368. The results of clustering using the method Ward are Cluster 1, areas that are quite safe, namely Aceh, Central Java, South Kalimantan, East Kalimantan, Lampung, East Nusa Tenggara, Papua, Riau, Central Sulawesi and North Sulawesi Provinces. Cluster 2 safe areas, namely Bali Province, Bangka Belitung Islands, Banten, Bengkulu, Gorontalo, West Papua, Jambi, West Kalimantan, Central Kalimantan, North Kalimantan, Riau Islands, North Maluku, Maluku, West Nusa Tenggara, West Sulawesi, Southeast Sulawesi, and DI Yogyakarta. Cluster 3 vulnerable areas, namely DKI Jakarta and North Sumatra Provinces. And cluster 4 areas are quite vulnerable, namely West Java, East Java, South Sulawesi, West Sumatra, and South Sumatra Provinces.*

**Keywords :** *Ward*, *K-Medoids*, Crime.

## PENDAHULUAN

Kriminalitas atau kejahatan adalah suatu perbuatan yang dapat mengakibatkan timbulnya masalah-masalah dan keresahan bagi kehidupan masyarakat. Berdasarkan hasil survey tingkat kriminalitas di Indonesia menduduki peringkat kelima di ASEAN (Asia, 2018). Semakin tinggi angka kriminalitas, menunjukkan semakin banyak tindak kejahatan pada masyarakat yang merupakan indikasi bahwa masyarakat merasa semakin tidak aman (Taram, 2019).

Menurut Abraham Maslow dalam teori hierarki kebutuhan manusia (Maslow, 1943), rasa aman berada pada tingkatan yang kedua di bawah kebutuhan dasar manusia seperti sandang, pangan, dan papan. Hal ini menunjukkan bahwa rasa aman merupakan kebutuhan manusia yang sangat penting. Rasa aman (*security*) merupakan salah satu hak asasi yang harus diperoleh setiap orang seperti yang disebutkan dalam UUD Republik Indonesia 1945 Pasal 28G ayat 1 yaitu :”Setiap orang berhak atas perlindungan diri pribadi, keluarga, kehormatan, martabat, dan harta benda yang di bawah kekuasaannya, serta berhak atas rasa aman dan perlindungan dari ancaman ketakutan untuk berbuat atau tidak berbuat sesuatu yang merupakan hak asasi”.

Jumlah kejadian kejahatan atau tindak kriminalitas di Indonesia pada tahun 2016-2018 cenderung fluktuatif. Menurut Badan Pusat Statistik, Jumlah kejadian kejahatan pada tahun 2016 sebanyak 357.197 kejadian, menurun menjadi sebanyak 336.652 kejadian pada tahun 2017 dan menurun pada tahun 2018 menjadi 294.28 kejadian. Dari segi jumlah kejahatan untuk provinsi/polda selama tahun 2018 polda Metro Jaya mencatat jumlah kejahatan terbanyak yaitu 34.655 kejadian, disusul oleh polda Sumatera Utara sebanyak 32.922 kejadian, dan polda Jawa Timur sebanyak 26.295 kejadian. Sementara itu, polda Kalimantan Utara, Maluku Utara, dan Sulawesi Tenggara merupakan tiga wilayah dengan jumlah kejahatan paling sedikit dengan jumlah kejadian kejahatan berturut-turut sebanyak 396, 722, dan 1.263 kejadian.

Indikator yang biasa digunakan untuk mengukur kejahatan adalah angka jumlah kejahatan (*crime total*), angka kejahatan per 100.000 penduduk (*crime rate*), dan selang waktu terjadinya suatu tindak kejahatan (*crime clock*). Dari tahun 2000 hingga 2016, resiko masyarakat terkena tindak kriminalitas makin tinggi dan selang waktu terjadi tindak kriminalitas itu makin singkat.

Pada tahun 2000 *crime clock* menunjukkan setiap 3 menit 2 detik terjadi satu tindak kriminal di Indonesia. Selang waktunya makin rapat di tahun 2016, yakni menjadi setiap 1 menit 28 detik terjadi kejahatan di Indonesia (Statistik, 2019).

Sejalan dengan *crime total*, tingkat resiko terkena tindak kejahatan (*crime rate*) setiap 100.000 penduduk mengalami penurunan pada tahun 2018 sebesar 113. Angka ini menurun dari 140 pada tahun 2016 dan 129 pada tahun 2017. Sebagai perbandingan, *crime rate* Malaysia (*Department of Statistics Malaysia*) pada tahun 2015 sebesar 370, tahun 2016 sebesar 355, dan tahun 2017 sebesar 309. *Crime rate* merupakan angka yang dapat menunjukkan tingkat kerawanan suatu kejahatan pada suatu kota tertentu dalam waktu tertentu. Semakin tinggi angka *crime rate* maka tingkat kerawanan akan kejahatan suatu daerah semakin tinggi pula, dan sebaliknya (Statistik, 2019).

Badan Pusat Statistik menunjukkan bahwa jumlah kasus kriminalitas di Indonesia selalu meningkat dari tahun ke tahun. Dengan banyaknya kasus kriminalitas di Indonesia perlu adanya pengklasteran daerah-daerah rawan kriminalitas di Indonesia, agar masyarakat waspada terhadap kejahatan dan pemerintah lebih memperhatikan dan memperketat sistem keamanan untuk daerah tersebut.

Analisis kluster mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Analisis kluster mengklasifikasi objek sehingga objek-objek yang paling dekat kesamaannya dengan objek lain berada dalam kluster yang sama (Ediyanto, 2013).

Metode *Ward (Wards Method)* adalah salah satu metode analisis hierarki dimana metode ini menggunakan *error sum of square (ESS)* sebagai pertimbangan dalam menggabungkan objek ke dalam kluster sehingga variansi di dalam kluster sekecil mungkin (Gudono, 2014). Dua obyek akan digabungkan jika mempunyai fungsi obyektif terkecil diantara kemungkinan yang ada. Metode *ward* merupakan suatu metode pembentukan kluster yang didasari oleh hilangnya informasi akibat penggabungan obyek menjadi kluster.

Metode *Ward (Wards Method)* adalah salah satu metode analisis hierarki dimana metode ini menggunakan *error sum of square (ESS)* sebagai pertimbangan dalam menggabungkan objek ke dalam kluster sehingga variansi di dalam kluster sekecil mungkin (Gudono, 2014). Dua obyek akan digabungkan jika mempunyai fungsi obyektif

terkecil diantara kemungkinan yang ada. Metode *ward* merupakan suatu metode pembentukan klaster yang didasari oleh hilangnya informasi akibat penggabungan obyek menjadi klaster.

Berdasarkan Penelitian terdahulu, penelitian dengan membandingkan hasil pengelompokan menggunakan metode *ward* dan *k-means* adalah penelitian yang dilakukan oleh Fiqih Satria dan RZ. Abdul Aziz (2016) yang membandingkan hasil kedua metode tersebut dalam menentukan *cluster* data mahasiswa pemohon beasiswa. Hasil dari penelitian ini menunjukkan bahwa metode *ward* menghasilkan nilai rasio simpangan baku sebesar 0,5346668% lebih kecil jika dibandingkan dengan nilai rasio simpangan baku pada metode *k-means* yaitu 0,831525302%. Dan metode *ward* merupakan metode yang lebih baik dibandingkan metode *k-means*.

Penelitian selanjutnya dilakukan oleh Farida Zulfa (2019) yang bertujuan untuk mengetahui hasil *cluster* yang terbentuk dan perbandingan hasil kinerja dengan menggunakan algoritma *k-means* dan *k-medoids* untuk pengklasteran provinsi di Indonesia berdasarkan indikator pendidikan. Dari perbandingan kedua metode tersebut, metode *k-medoids* adalah metode terbaik karena memiliki nilai simpangan baku dalam klaster yaitu 0,236952 dan nilai simpangan baku antar klaster yaitu 0,289465 sehingga nilai rasio simpangan baku yaitu 81,85867.

Penelitian lain dilakukan oleh Yusuf dan Novian (2014) yang bertujuan untuk merancang dan mengimplementasikan sistem pengklasteran ikan khususnya pada ikan cupang dengan menggunakan fitur warna, bentuk, dan tekstur. Penelitian ini menghasilkan *clustering* data ikan cupang menggunakan metode *k-means* dan *k-medoids* untuk temu citra kembali yang mampu mengklasterisasi data citra *database* dengan jumlah yang banyak dengan baik. Penelitian ini juga membuktikan bahwa metode *k-medoids* mendapatkan *clustering* yang lebih akurat dengan nilai *running time* lebih cepat dibandingkan metode *k-means*.

Penelitian ini akan melakukan analisis pengklasteran pada kasus kriminalitas di Indonesia. Berdasarkan (Statistik, 2019), kejadian kejahatan kriminalitas diklasifikasikan menjadi kejahatan terhadap nyawa, kejahatan terhadap fisik/badan, kejahatan terhadap kesusilaan, kejahatan terhadap kemerdekaan orang, kejahatan terhadap hak milik/barang dengan penggunaan kekerasan, kejahatan terhadap hak milik/barang tanpa

penggunaan kekerasan, kejahatan terkait narkoba, dan kejahatan terkait penipuan, penggelapan dan korupsi. Berdasarkan kejadian-kejadian kejahatan kriminalitas tersebut maka dilakukan analisis klastering menggunakan metode *ward* dan *k-medoids* untuk mengetahui pengklasteran kasus kriminalitas yang ada di setiap Provinsi di Indonesia. Setelah diperoleh hasil pengklasteran, kemudian dilakukan perhitungan rasio simpangan baku. Rasio yang diperoleh dari perbandingan simpangan baku dalam klaster dengan simpangan baku antar klaster berguna untuk mengetahui metode mana yang terbaik dalam mengelompokkan Provinsi di Indonesia berdasarkan kriminalitas.

Tujuan penelitian ini yaitu untuk mengetahui gambaran umum tingkat kriminalitas di Indonesia, untuk mengetahui hasil pengelompokan provinsi di Indonesia berdasarkan kriminalitas dengan menggunakan metode *ward* dan *k-medoids*. Dan untuk mengetahui hasil pengelompokan terbaik antara metode *ward* dan *k-medoids*.

## TINJAUAN PUSTAKA

### 1. Kriminalitas

Menurut BPS (Badan Pusat Statistik), Tindak kejahatan/kriminalitas atau pelanggaran merupakan perbuatan seseorang yang dapat diancam hukuman berdasarkan kitab Undang-Undang Hukum Pidana (KUHP) atau undang-undang serta peraturan lainnya yang berlaku di Indonesia.

### 2. Kejahatan terhadap Nyawa

Menurut BPS (Badan Pusat Statistik), Kejahatan terhadap nyawa (pembunuhan/*homicide*) merupakan salah satu kejahatan yang paling tinggi hierarkinya dalam klasifikasi kejahatan internasional, selain itu dari sisi hukuman juga yang paling berat hukumannya dalam KUHP Indonesia.

### 3. Kejahatan terhadap Fisik/Badan

Menurut BPS (Badan Pusat Statistik), kejahatan terhadap fisik termasuk di dalamnya adalah penganiayaan ringan, penganiayaan berat, dan kekerasan dalam rumah tangga. Penganiayaan adalah perbuatan dengan sengaja merusak kesehatan fisik orang lain baik menimbulkan penyakit (luka/cacat/sakit) atau halangan untuk menjalankan pekerjaan sehari-hari (luka berat) maupun tidak. Pemukulan, penamparan, pengeroyokan, termasuk kategori penganiayaan.

### 4. Kejahatan terhadap Kesusilaan

Menurut BPS (Badan Pusat Statistik), dalam klasifikasi kejahatan terhadap kesusilaan terdiri

dari jenis kejahatan perkosaan dan pencabulan. Pelecehan seksual adalah perbuatan yang menyerang kehormatan kesusilaan (cubitan, colekan, tepukan, sentuhan di bagian tubuh tertentu atau gerakan) maupun perbuatan cabul yang dilakukan secara sepihak dan tidak diharapkan oleh korban. Perbuatan tersebut dapat dilakukan dengan kekerasan fisik, ancaman kekerasan, maupun tidak. Perkosaan (perbuatan yang dilakukan dengan kekerasan atau ancaman kekerasan, memaksa korban untuk bersetubuh dengannya) termasuk kategori pelecehan seksual.

#### **5. Kejahatan terhadap Kemerdekaan Orang**

Menurut BPS (Badan Pusat Statistik), dalam klasifikasi kejahatan terhadap kemerdekaan orang terdiri dari jenis kejahatan penculikan dan mempekerjakan anak dibawah umur. Perniagaan budak belian adalah perniagaan yang barang dagangannya terdiri dari orang-orang yang akan dipergunakan untuk dijadikan budak atau hamba belian.

#### **6. Kejahatan terhadap Hak Milik/Barang dengan Penggunaan Kekerasan**

Menurut BPS (Badan Pusat Statistik), dalam klasifikasi kejahatan ini yang termasuk kejahatan terhadap hak/milik dengan penggunaan kekerasan adalah pencurian dengan kekerasan menggunakan senjata api (senpi) dan pencurian dengan kekerasan menggunakan senjata tajam (sajam).

#### **7. Kejahatan terhadap Hak Milik/Barang tanpa Penggunaan Kekerasan**

Menurut BPS (Badan Pusat Statistik), dalam klasifikasi kejahatan ini yang termasuk adalah pencurian, pencurian dengan pemberatan, pencurian kendaraan bermotor, pengrusakan/penghancuran barang, pembakaran dengan sengaja, dan penadahan.

#### **8. Kejahatan terkait Narkotika**

Menurut BPS (Badan Pusat Statistik), dalam klasifikasi kejahatan ini adalah yang menggunakan narkotika dan psikotropika. Kejahatan narkotika merupakan suatu kejadian transnasional yang saat ini menjadi suatu kejahatan lintas negara yang terjadi di Indonesia. Pada 2018, jenis kejahatan yang termasuk kejahatan luar biasa (*extraordinary crime*) ini memiliki pola yang berbeda dengan kejahatan lain.

#### **9. Kejahatan terkait Penipuan, Penggelapan, dan Korupsi**

Menurut BPS (Badan Pusat Statistik), dalam klasifikasi kejahatan ini yang termasuk adalah

penipuan, penggelapan, dan korupsi. Tindak pidana penggelapan diatur dalam pasal 372 KUHP yang berbunyi “Barang siapa dengan sengaja menguasai secara melawan hukum sesuatu benda yang seharusnya atau sebagian merupakan kepunyaan orang lain yang berada padanya bukan karena kejahatan, karena bersalah melakukan penggelapan, dipidana dengan pidana penjara selama-lamanya 4 (empat) tahun atau dengan pidana denda setinggi-tingginya 900 (sembilan ratus) rupiah”.

#### **10. Analisis Klaster**

*Cluster* atau ‘klaster’ dapat diartikan ‘klaster;’ dengan demikian, pada dasarnya analisis klaster akan menghasilkan sejumlah klaster (klaster) (Santoso, 2014 : 125). Analisis klaster yaitu analisis untuk mengelompokkan elemen yang mirip sebagai objek penelitian menjadi klaster yang berbeda dan *mutually exclusive* (Supranto, 2004 : 26).

Tujuan utama analisis klaster ialah mengklasifikasi objek (kasus/elemen) seperti orang, produk (barang), toko, perusahaan kedalam klaster-klaster yang relatif homogen didasarkan pada suatu set variabel yang dipertimbangkan untuk diteliti (Supranto, 2004 : 141).

#### **11. Proses Analisis Klaster**

Menurut (Santoso : 2014), proses analisis klaster adalah sebagai berikut :

##### **1. Proses Standarisasi**

Bila terdapat data (variabel) yang mempunyai perbedaan ukuran satuan yang besar perlu dilakukan proses standarisasi. Perbedaan satuan yang mencolok akan menyebabkan bias dalam analisis klaster, sehingga data asli harus ditransformasi (standarisasi) sebelum bisa dianalisis. Transformasi dilakukan terhadap variabel yang relevan ke dalam bentuk z skor, seperti berikut :

$$Z = \frac{x - \bar{x}}{\sigma}$$

Keterangan :

$x$  = nilai data

$\bar{x}$  = nilai rata-rata

$\sigma$  = standar deviasi

##### **2. Mengukur Kesamaan Antar Objek (*Similarity*)**

Konsep kesamaan adalah hal yang fundamental dalam analisis klaster. Kesamaan antar objek merupakan ukuran korespondensi antar objek. Ada tiga metode yang dapat diterapkan, yaitu ukuran korelasi, ukuran jarak dan ukuran asosiasi.

### 3. Membuat kluster (*Clustering*)

Ada dua proses penting yaitu algoritma kluster dalam pembentukan kluster dan menentukan jumlah kluster yang akan dibentuk. Keduanya mempunyai implikasi substansial tidak hanya pada hasil yang diperoleh tetapi juga pada interpretasi yang akan dilakukan terhadap hasil tersebut (Supranto, 2004).

### 4. Interpretasi

Setelah kluster terbentuk, langkah selanjutnya melakukan interpretasi. Tahap interpretasi meliputi pengujian tiap kluster dalam term untuk menamai dan menandai dengan suatu label secara akurat dapat menjelaskan kealamian kluster. Proses ini dimulai dengan suatu ukuran yang sering digunakan yaitu centroid cluster.

### 5. Validasi dan *Profiling* Kluster

Proses validasi bertujuan untuk menjamin bahwa solusi yang dihasilkan dari analisis kluster dapat mewakili populasi dan dapat digeneralisasi untuk objek lain. Pendekatan ini membandingkan solusi kluster dan menilai korespondensi hasil. Terkadang tidak dapat dipraktekkan karena adanya kendala waktu dan biaya atau ketidakterediaan objek untuk analisis kluster ganda (Supranto, 2004).

### 12. Analisis Komponen Utama

Jika terdapat variabel yang terjadi multikolinearitas, maka salah satu solusi untuk menangani masalah tersebut adalah dengan analisis komponen utama. Analisis komponen utama adalah suatu analisis yang bertujuan untuk mentransformasikan  $p$  variabel asal yang masih berkorelasi satu dengan yang lain menjadi satu set variabel baru yang tidak berkorelasi lagi.

Pembentukan *Principal Component* (PC) merupakan suatu kombinasi linear dari variabel-variabel asal. Pembentukan *Principal Component* (PC) terdapat dua cara yaitu dengan matriks kovarian atau matriks korelasi (Johnson dan Wichern, 2007).

### 13. *Ward's Error Sum of Square Method*

Metode *Ward's* ditemukan oleh Ward (1963) tidak menghitung jarak antar kluster atau objek, namun metode ini membentuk kluster-kluster dengan memaksimalkan kehomogenan dalam kluster. Metode *ward* merupakan suatu metode pembentukan kluster yang didasari oleh hilangnya

informasi akibat penggabungan obyek menjadi kluster. Hal ini diukur dengan menggunakan jumlah total dari deviasi kuadrat pada *mean* kluster untuk setiap pengamatan. *Error sum of squares* (SSE) digunakan sebagai fungsi obyektif. Dua obyek akan digabungkan jika mempunyai fungsi obyektif terkecil diantara kemungkinan yang ada.

$$SSE = \sum_{j=1}^p \left( \sum_{i=1}^n x_{ij}^2 - \frac{1}{n} \left( \sum_{i=1}^n x_{ij} \right)^2 \right)$$

Dimana :

$x_{ij}$  : nilai untuk obyek ke-I pada kluster ke-j

$p$  : banyaknya variabel yang diukur

$n$  : banyaknya obyek dalam kluster yang terbentuk

Langkah penyelesaian metode *Ward* yaitu :

1. Dimulai dengan memperhatikan  $N$  kluster yang mempunyai satu responden per kluster (semua responden dianggap sebagai kluster). Pada tahap pertama ini SSE bernilai nol.
2. Kluster pertama dibentuk dengan memilih dua dari  $N$  kluster yang memiliki nilai SSE terkecil.
3.  $N-1$  kumpulan kluster kemudian diperhatikan kembali untuk menentukan dua dari kluster ini yang bisa meminimumkan keheterogenan. Dengan demikian  $N$  kluster secara sistematis dikurangi  $N-1$ .
4. Mengulangi langkah 3 dan 4 sampai diperoleh satu kluster atau semua responden bergabung menjadi satu kluster.

### 14. *K-Medoids Cluster*

*K-Medoids Clustering* juga dikenal sebagai *Partitioning Around Medoids (PAM)* adalah varian dari metode *k-means*. Hal ini didasarkan pada penggunaan *medoids* bukan dari pengamatan *mean* yang dimiliki oleh setiap kluster, dengan tujuan mengurangi sensitivitas dari partisi sehubungan dengan nilai ekstrim yang ada dalam dataset (Vercellis, 2009). *K-Medoids Clustering* hadir untuk mengatasi kelemahan *k-means clustering* yang sensitif terhadap *outlier* karena suatu objek dengan suatu nilai yang besar mungkin secara substansial menyimpang dari distribusi data (Han dan Kamber, 2006).

Menurut Han dan Kamber (2006), tahapan *k-medoids clustering* adalah sebagai berikut.

1. Secara acak pilih  $k$  objek pada sekumpulan  $n$  objek sebagai *medoids*.
2. Ulangi.
3. Tempatkan objek non *medoids* ke dalam kluster yang paling dekat dengan *medoids*.

4. Secara acak pilih  $O_{random}$  (sebuah objek *non medoids*).
5. Hitung total *cost*,  $S$ , dari pertukaran *medoids*  $O_j$  dengan  $O_{random}$ .  
 $S = \text{total cost baru} - \text{total cost lama}$   
 Dengan :  
 $S$  : Selisih  
 Total *cost* baru : Total *cost* baru untuk *non medoids*  
 Total *cost* lama : Total *cost* lama untuk *medoids*
6. Jika  $S < 0$  maka tukar  $O_j$  dengan  $O_{random}$ , untuk membentuk sekumpulan  $k$  objek baru sebagai *medoids*.
7. Hingga tidak ada perubahan.

### 15. Perbandingan Model Terbaik

Untuk mengetahui metode mana yang mempunyai kinerja terbaik, dapat digunakan rata-rata simpangan baku dalam kluster ( $\sigma_w$ ) dan simpangan baku antar kluster ( $\sigma_B$ ) (Bunkers, dkk, 1996).

Rumus rata-rata simpangan baku dalam kluster yaitu :

$$(\sigma_w) = K^{-1} \sum_{k=1}^k \sigma_k$$

dengan :

$K$  : Banyaknya kluster yang terbentuk

$\sigma_k$  : Simpangan baku kluster ke- $k$

Rumus simpangan baku kluster ke- $k$  yaitu:

$$\sigma_k = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \mu_k)^2}$$

dengan :

$N$  : Jumlah anggota dari setiap kluster

$\mu_k$  : Rata-rata kluster ke- $k$

$X_i$  : Anggota kluster dari  $i=1,2,\dots,N$

Rumus simpangan baku antar kluster  $\sigma_B$  yaitu:

$$\sigma_B = [(K)^{-1} \sum_{k=1}^K (\mu_k - \mu)^2]^{1/2}$$

dengan :

$\mu_k$  : Rata-rata kluster ke- $k$

$\mu$  : Rata-rata keseluruhan kluster

Rumus rasio simpangan baku kluster  $\sigma$  yaitu :

$$\sigma = \frac{\sigma_w}{\sigma_B} \times 100\%$$

dengan :

$\sigma_w$ : simpangan baku dalam kluster

$\sigma_B$ : simpangan baku antar kluster

Metode yang mempunyai rasio terkecil merupakan metode terbaik. Kluster yang baik adalah kluster yang mempunyai homogenitas (kesamaan) yang tinggi antar anggota dalam satu kluster (*within cluster*) dan heterogenitas yang tinggi antar kluster yang satu dengan kluster yang lain (*between cluster*) (Santoso, 2007).

## METODE PENELITIAN

### 1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh melalui publikasi lembaga pemerintahan resmi yaitu Badan Pusat Statistik Provinsi Sumatera Utara periode tahun 2018. Unit observasi dalam penelitian ini adalah Provinsi-Provinsi di seluruh Indonesia.

### 2. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini ada 8 kejahatan terhadap nyawa, kejahatan terhadap fisik/badan, kejahatan terhadap kesusilaan, kejahatan terhadap kemerdekaan orang, kejahatan terhadap hak milik/barang dengan penggunaan kekerasan, kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan, kejahatan terkait narkoba, dan kejahatan terkait penipuan, penggelapan dan korupsi dari 34 Provinsi di Indonesia.

### 3. Analisis Data

Tahapan analisis yang digunakan dalam penelitian ini adalah :

1. Melakukan analisis statistika deskriptif
2. Uji asumsi Multikolinearitas untuk mengetahui ukuran kemiripan apa yang dapat digunakan. Jika terdapat variabel yang terjadi multikolinearitas, maka salah satu solusi untuk menangani masalah tersebut adalah dengan analisis komponen utama.
3. Pengklasteran dengan metode *Ward* dan *K-Medoids*.
4. Membandingkan metode *ward* dan metode *k-medoids* dengan melihat rasio simpangan baku terkecil untuk mendapatkan metode terbaik.

## HASIL DAN PEMBAHASAN

### 1. Deskriptif Data

Gambaran umum data penelitian disajikan pada tabel 4.1. yang memuat nilai minimum, mean, maksimum, dan standar deviasi dari kedelapan variabel yang digunakan.

Variabel	Minimum	Mean	Maximum	Standar deviasi
X1	2	30.12	111	38.849
X2	39	1163.7	5240	1206.503
X3	26	154.6	512	116.623
X4	0	74.85	440	92.945
X5	2	213.24	702	204.449
X6	125	2749	10818	2531.216
X7	6	1164.4	8715	1803.898
X8	45	1290	5526	1442.344

N=34

Dilihat dari tabel diatas dapat diketahui nilai masing-masing variabel, untuk variabel kejahatan terhadap nyawa (X1) memiliki nilai minimum 2, dimana nilai ini merupakan jumlah kejahatan terhadap nyawa paling sedikit di provinsi DI Yogyakarta, mean sebesar 30.12 nilai ini merupakan rata-rata jumlah kejadian kejahatan terhadap nyawa, nilai maksimum 111 yang menggambarkan jumlah kejahatan terhadap nyawa terbanyak di Provinsi Sumatera Selatan, dan standar deviasi sebesar 38.849. Variabel kejahatan terhadap fisik/badan (X2) memiliki nilai minimum 39 dimana nilai ini merupakan jumlah kejahatan terhadap fisik/badan paling sedikit di provinsi Kalimantan Utara, mean 1163.7 nilai ini merupakan rata-rata jumlah kejadian kejahatan terhadap fisik/badan, nilai maksimum 5240 yang menggambarkan jumlah kejahatan terhadap fisik/badan terbanyak di Provinsi Sumatera Utara, dan standar deviasi sebesar 1206.503. Variabel kejahatan terhadap kesusilaan (X3) memiliki nilai minimum 26 dimana nilai ini merupakan jumlah kejahatan terhadap kesusilaan paling sedikit di provinsi Kalimantan Utara, mean 154.6 nilai ini merupakan rata-rata jumlah kejadian kejahatan terhadap kesusilaan, nilai maksimum 512 yang menggambarkan jumlah kejahatan terhadap kesusilaan terbanyak di Provinsi Jawa Barat, dan standar deviasi sebesar 116.623. Variabel kejahatan terhadap kemerdekaan orang (X4) memiliki nilai minimum 0 dimana nilai ini artinya tidak ada jumlah kejahatan terhadap kemerdekaan orang di provinsi Kalimantan Timur, mean 74.85 nilai ini merupakan rata-rata jumlah kejadian kejahatan terhadap kemerdekaan orang, nilai

maksimum 440 yang menggambarkan jumlah kejahatan terhadap kemerdekaan orang terbanyak di Provinsi Sulawesi Selatan, dan standar deviasi sebesar 92.945. Variabel kejahatan terhadap hak milik/barang dengan penggunaan kekerasan (X5) memiliki nilai minimum 2 dimana nilai ini merupakan jumlah kejahatan terhadap hak milik/barang dengan penggunaan kekerasan paling sedikit di provinsi Maluku Utara, mean 213.24 nilai ini merupakan rata-rata jumlah kejadian kejahatan terhadap hak milik/barang dengan penggunaan kekerasan, nilai maksimum 702 yang menggambarkan jumlah kejahatan terhadap hak milik/barang dengan penggunaan kekerasan orang terbanyak di Provinsi Sumatera Utara, dan standar deviasi sebesar 204.449. Variabel kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan (X6) memiliki nilai minimum 125 dimana nilai ini merupakan jumlah kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan paling sedikit di provinsi Maluku Utara, mean 2749 nilai ini merupakan rata-rata jumlah kejadian kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan, nilai maksimum 10818 yang menggambarkan jumlah kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan terbanyak di Provinsi Sumatera Utara, dan standar deviasi sebesar 2531.216. Variabel kejahatan terkait narkoba (X7) memiliki nilai minimum 6 dimana nilai ini merupakan jumlah kejahatan terkait narkoba paling sedikit di provinsi Nusa Tenggara Timur, mean 1164.4 nilai ini merupakan rata-rata jumlah kejadian kejahatan terkait narkoba, nilai maksimum 8715 yang menggambarkan jumlah kejahatan terkait narkoba terbanyak di Provinsi DKI Jakarta, dan standar deviasi sebesar 1803.898. Dan variabel kejahatan terkait penipuan, penggelapan dan korupsi (X8) memiliki nilai minimum 45 dimana nilai ini merupakan jumlah kejahatan terkait penipuan, penggelapan dan korupsi paling sedikit di provinsi Maluku Utara, mean 1290 nilai ini merupakan rata-rata jumlah kejadian kejahatan terkait penipuan, penggelapan dan korupsi, nilai maksimum 5526 yang menggambarkan jumlah kejahatan terkait penipuan, penggelapan dan korupsi terbanyak di Provinsi DKI Jakarta, dan standar deviasi sebesar 1442.344.

## 2. Uji Asumsi Multikolinearitas

Uji asumsi multikolinearitas bertujuan untuk mengetahui apakah ada korelasi yang tinggi antar variabel. Jika tidak terjadi multikolinearitas maka data dapat langsung dianalisis, tetapi jika terjadi multikolinearitas maka harus dilakukan penanganan data terlebih dahulu sebelum data dianalisis. Pengujian untuk memastikan ada tidaknya multikolinearitas yaitu dengan menggunakan nilai VIF > 10. Selain itu, bisa dengan melihat nilai dari matrik korelasi antar variabel. Apabila nilai korelasi antar variabel lebih dari 0,70 maka terjadi multikolinearitas (Puspitasari, 2016).

Variabel	X1	X2	X3	X4	X5	X6	X7	X8
X1	1	<b>0,75</b>	0,47	0,38	0,63	0,67	0,61	<b>0,72</b>
X2	<b>0,75</b>	1	0,70	0,51	0,40	<b>0,87</b>	0,46	<b>0,75</b>
X3	0,47	0,70	1	0,32	0,32	<b>0,71</b>	0,23	0,65
X4	0,38	0,51	0,32	1	0,06	0,53	-	0,30
X5	0,63	0,40	0,32	0,06	1	0,60	0,62	0,61
X6	0,67	<b>0,87</b>	<b>0,71</b>	0,53		1	0,51	<b>0,81</b>
X7	0,61	0,46	0,23	-	0,60	0,62	0,51	1
X8	<b>0,72</b>	<b>0,75</b>	0,65	0,30	0,61	<b>0,81</b>	<b>0,78</b>	1

Dari tabel diatas dapat dilihat bahwa ada nilai korelasi yang lebih besar dari 0.70 yaitu korelasi antar variabel X1 dengan X2 yaitu 0.75, variabel X1 dengan X8 yaitu 0.72, variabel X2 dengan X6 yaitu 0.87, variabel X3 dengan X6 yaitu 0.71, variabel X8 dengan X2 yaitu 0.75, variabel X8 dengan X6 yaitu 0.81, dan variabel X8 dengan X7 yaitu 0.78. Maka dapat dikatakan terjadinya multikolinearitas. Karena terjadinya multikolinearitas, maka perlu dilakukan upaya perbaikan yaitu dengan reduksi variabel melalui *Principal Component Analysis* (PCA) atau Analisis Komponen Utama (AKU) untuk memperoleh variabel baru yang akan digunakan dalam analisis klaster.

## 3. Analisis Komponen Utama

Analisis komponen utama atau *participal component analysis* (PCA) adalah salah satu solusi untuk menangani data yang multikolinearitas.

Komponen Utama ke i	Nilai Eigen	Proporsi Varian	Proporsi Kumulatif
KU 1	1.05	0.8909	0.8909
KU 2	2.19	0.1036	<b>0.9944</b>
KU 3	2.87	0.00367	0.99812
KU 4	2.41	0.00155	0.99968
KU 5	1.69	0.00024	0.99992
KU 6	6.13	0.00008	1
KU 7	2.87	0	1
KU 8	1.53	0	1

Berdasarkan tabel di atas, didapatkan nilai proporsi kumulatif sebesar 0,9944, artinya sebesar

99,44% varians dari 8 variabel kriminalitas dapat dijelaskan oleh 2 skor komponen utama.

Selanjutnya untuk menentukan skor komponen utama menggunakan nilai *component loadings*.

Variabel	Comp 1	Comp2
X1	0.3712175	-0.1211107
X2	0.3691434	0.1623465
X3	0.3613636	0.2791975
X4	0.3123218	0.6050489
X5	0.3603211	-0.2641719
X6	0.3708881	0.1431901
X7	0.3083500	-0.6226072
X8	0.3678765	-0.1923807

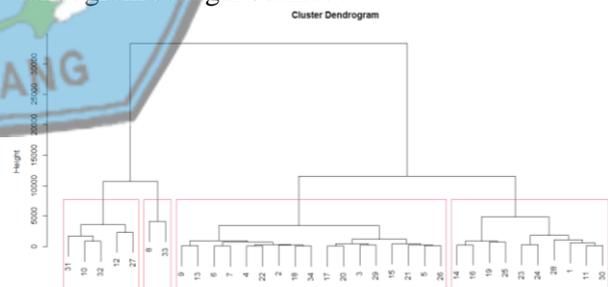
Berdasarkan tabel di atas, dapat diketahui nilai *component loadings* kedua komponen utama sehingga diperoleh persamaan :

$$KU_1 = 0.3712175X_1 + 0.3691434X_2 + 0.3613636X_3 + 0.3123218X_4 + 0.3603211X_5 + 0.3708881X_6 + 0.3083500X_7 + 0.3678765X_8$$

$$KU_2 = -0.1211107 + 0.1623465X_2 + 0.2791975X_3 + 0.6050489X_4 - 0.2641719X_5 + 0.1431901X_6 - 0.6226072X_7 - 0.1923807X_8$$

## 4. Proses Pengklasteran Ward

Analisis ini menggunakan aplikasi R dan diperoleh hasil Pengklasteran dalam bentuk dendrogram sebagai berikut :



Berdasarkan Gambar di atas, terlihat bahwa masing-masing objek mengelompok secara hierarki sehingga setiap objek bisa diperkirakan masuk kedalam klaster yang mana berdasarkan banyaknya klaster yang akan dibentuk.



Berdasarkan gambar di atas, didapatkan hasil Pengklasteran dengan metode *ward* yaitu : Kluster 1 ditandai dengan warna biru muda yaitu Provinsi Aceh, Jawa Tengah, Kalimantan Selatan, Kalimantan Timur, Lampung, Nusa Tenggara Timur, Papua, Riau, Sulawesi Tengah, dan Sulawesi Utara. Kluster 2 yaitu ditandai dengan warna biru tua yaitu Provinsi Bali, Kepulauan Bangka Belitung, Banten, Bengkulu, Gorontalo, Papua Barat, Jambi, Kalimantan Barat, Kalimantan Tengah, Kalimantan Utara, Kepulauan Riau, Maluku Utara, Maluku, Nusa Tenggara Barat, Sulawesi Barat, Sulawesi Tenggara, dan DI Yogyakarta. Kluster 3 yaitu ditandai dengan warna hijau muda yaitu Provinsi DKI Jakarta dan Sumatera Utara. Dan kluster 4 ditandai dengan warna hijau tua yaitu Provinsi Jawa Barat, Jawa Timur, Sulawesi Selatan, Sumatera Barat, dan Sumatera Selatan.

Adapun karakteristik dari masing-masing kluster yaitu :

Kluster	Variabel	Rata-rata
1	kejahatan terhadap nyawa	32,5
	kejahatan terhadap fisik/badan	1187,6
	kejahatan terhadap kesusilaan	165,3
	kejahatan terhadap hak milik/barang dengan penggunaan kekerasan	254,6
	kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan	2743,6
	kejahatan terkait narkoba	1183,6
	kejahatan terkait penipuan, penggelapan dan korupsi	1022,9
2	kejahatan terhadap nyawa	12,53
	kejahatan terhadap fisik/badan	461,65
	kejahatan terhadap kesusilaan	85,94
	kejahatan terhadap kemerdekaan orang	51,41
	kejahatan terhadap hak milik/barang dengan penggunaan kekerasan	85,59
	kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan	1126,65
	kejahatan terkait narkoba	290,88
3	kejahatan terhadap penipuan, penggelapan dan korupsi	437,24
	kejahatan terhadap nyawa	91
	kejahatan terhadap fisik/badan	3623
	kejahatan terhadap hak milik/barang dengan penggunaan kekerasan	553,5

kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan	7583,5
kejahatan terkait narkoba	7488,5
kejahatan terkait penipuan, penggelapan dan korupsi	5103
kejahatan terhadap nyawa	60,8
kejahatan terhadap fisik/badan	2519,4
kejahatan terhadap hak milik/barang dengan penggunaan kekerasan	428,4
4 kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan	6340,2
kejahatan terkait narkoba	1566
kejahatan terkait penipuan, penggelapan dan korupsi	3196,8

1. Kluster 1 yaitu memiliki beberapa karakteristik yaitu kejahatan terhadap nyawa memiliki nilai rata-rata 32.5, kejahatan terhadap fisik/badan memiliki nilai rata-rata 1187.6, kejahatan terhadap kesusilaan memiliki nilai rata-rata 165.3, kejahatan terhadap hak milik/barang dengan penggunaan kekerasan memiliki nilai rata-rata 254.6, kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan memiliki nilai rata-rata 2743.6, kejahatan terkait narkoba memiliki nilai rata-rata 1183.6, dan kejahatan terkait penipuan, penggelapan dan korupsi memiliki nilai rata-rata 1022.9. Dimana dapat diartikan bahwa pada kluster 1 terdapat 7 variabel dengan nilai rata-rata sedang. Oleh karena itu, kluster 1 yaitu masuk ke dalam kategori daerah kriminalitas cukup aman karena memiliki nilai rata-rata sedang.
2. Kluster 2 memiliki beberapa karakteristik yaitu kejahatan terhadap nyawa, kejahatan terhadap fisik/badan, kejahatan terhadap kesusilaan, kejahatan terhadap kemerdekaan orang, kejahatan terhadap hak milik/barang dengan penggunaan kekerasan, kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan, kejahatan terkait narkoba, dan kejahatan terkait penipuan, penggelapan dan korupsi memiliki nilai rata-rata terendah dibandingkan dengan kluster lainnya. Dimana dapat diartikan bahwa pada kluster 2 semua variabel memiliki nilai dibawah rata-rata. Oleh karena itu, kluster 2 yaitu masuk ke dalam kategori daerah kriminalitas aman karena memiliki nilai di bawah rata-rata.
3. Kluster 3 memiliki beberapa karakteristik yaitu kejahatan terhadap nyawa, kejahatan terhadap fisik/badan, kejahatan terhadap hak milik/barang dengan penggunaan kekerasan, kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan, kejahatan terkait narkoba, dan kejahatan terkait penipuan,

penggelapan dan korupsi memiliki nilai rata-rata tertinggi dibandingkan dengan klaster lainnya. Dimana dapat diartikan bahwa pada klaster 3 semua variabel memiliki nilai diatas rata-rata. Oleh karena itu, klaster 3 yaitu masuk ke dalam kategori daerah kriminalitas rawan karena memiliki nilai di atas rata-rata.

4. Klaster 4 memiliki beberapa karakteristik yaitu kejahatan terhadap nyawa memiliki nilai rata-rata 60.8, kejahatan terhadap fisik/badan memiliki nilai rata-rata 2519.4, kejahatan terhadap hak milik/barang dengan penggunaan kekerasan memiliki nilai rata-rata 428.4, kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan memiliki nilai rata-rata 6340.2, kejahatan terkait narkoba, dan kejahatan terkait penipuan memiliki nilai rata-rata 1566, penggelapan dan korupsi memiliki nilai rata-rata 3196.8. Dimana dapat diartikan bahwa pada klaster 4 terdapat 6 variabel memiliki nilai rata-rata sedang tetapi lebih tinggi dari nilai rata-rata pada klaster 1. Oleh karena itu, klaster 4 yaitu masuk ke dalam kategori daerah kriminalitas cukup rawan.

### 5. Proses Pengklasteran K-Medoids

Adapun hasil pengklasteran dengan metode *k-medoids* yaitu:



Berdasarkan gambar di atas, didapatkan hasil pengklasteran dengan metode *k-medoids* yaitu : Klaster 1 ditandai dengan warna biru muda yaitu Provinsi Aceh, Jawa Tengah, Lampung, Nusa Tenggara Timur, Papua, Riau, Sulawesi Tengah, dan Sulawesi Utara. Klaster 2 yaitu ditandai dengan warna biru tua yaitu Provinsi Bali, Kepulauan Bangka Belitung, Banten, Bengkulu, Gorontalo, Papua Barat, Jambi, Kalimantan Barat, Kalimantan Selatan, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Kepulauan Riau, Maluku Utara, Maluku, Nusa Tenggara Barat, Sulawesi Barat, Sulawesi Tenggara, dan DI Yogyakarta. Klaster 3 yaitu ditandai dengan warna

hijau muda yaitu Provinsi DKI Jakarta dan Sumatera Utara. Dan klaster 4 ditandai dengan warna hijau tua yaitu Provinsi Jawa Barat, Jawa Timur, Sulawesi Selatan, Sumatera Barat, dan Sumatera Selatan.

Adapun karakteristik dari masing-masing klaster yaitu :

Klaster	Variabel	Rata-rata
1	kejahatan terhadap nyawa	33,25
	kejahatan terhadap fisik/badan	1369
	kejahatan terhadap kesusilaan	186.88
	kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan	3058.38
	kejahatan terkait narkoba	1009.75
	kejahatan terkait penipuan, penggelapan dan korupsi	1146
2	kejahatan terhadap nyawa	14.32
	kejahatan terhadap fisik/badan	461,68
	kejahatan terhadap kesusilaan	85,21
	kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan	1164,32
	kejahatan terkait narkoba	458.05
	kejahatan terkait penipuan, penggelapan dan korupsi	447.05
3	kejahatan terhadap fisik/badan	2190.50
	kejahatan terhadap kesusilaan	228.50
	kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan	3411.50
4	kejahatan terhadap fisik/badan	2519.40
	kejahatan terhadap kesusilaan	337.40
	kejahatan terhadap kemerdekaan orang	179.60
	kejahatan terhadap hak milik/barang dengan penggunaan kekerasan	428.40
	kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan	6340.20

1. Klaster 1 yaitu memiliki beberapa karakteristik yaitu kejahatan terhadap nyawa memiliki nilai rata-rata 32.25, kejahatan terhadap fisik/badan memiliki nilai rata-rata 1369, kejahatan terhadap kesusilaan memiliki nilai rata-rata 186.88, kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan memiliki nilai rata-rata 3058.38, kejahatan terkait narkoba memiliki nilai rata-rata 1009.75, dan kejahatan terkait penipuan, penggelapan dan korupsi memiliki nilai rata-rata 1146. Dimana dapat diartikan bahwa pada klaster 1 terdapat 6 variabel dengan nilai rata-rata sedang. Oleh karena itu, klaster 1 yaitu masuk ke dalam kategori daerah kriminalitas cukup aman karena memiliki nilai rata-rata sedang.

2. Klaster 2 memiliki beberapa karakteristik yaitu kejahatan terhadap nyawa, kejahatan terhadap fisik/badan, kejahatan terhadap kesusilaan, kejahatan terhadap hak milik/barang tanpa

penggunaan kekerasan, kejahatan terkait narkoba, dan kejahatan terkait penipuan, penggelapan dan korupsi memiliki nilai rata-rata terendah dibandingkan dengan klaster lainnya. Dimana dapat diartikan bahwa pada klaster 2 semua variabel memiliki nilai dibawah rata-rata. Oleh karena itu, klaster 2 yaitu masuk ke dalam kategori daerah kriminalitas aman karena memiliki nilai di bawah rata-rata.

3. Klaster 3 memiliki beberapa karakteristik yaitu kejahatan terhadap fisik/badan memiliki nilai rata-rata 2190.50, kejahatan terhadap kesusilaan memiliki nilai rata-rata 228.50, dan kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan memiliki nilai rata-rata 3411.50. Dimana dapat diartikan bahwa pada klaster 3 terdapat 3 variabel memiliki nilai rata-rata sedang tetapi lebih tinggi dari nilai rata-rata pada klaster 1. Oleh karena itu, klaster 3 yaitu masuk ke dalam kategori daerah kriminalitas cukup rawan.
4. Klaster 4 memiliki beberapa karakteristik yaitu kejahatan terhadap fisik/badan, kejahatan terhadap kesusilaan, kejahatan terhadap kemerdekaan orang, kejahatan terhadap hak milik/barang dengan penggunaan kekerasan, kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan memiliki nilai rata-rata tertinggi dibandingkan dengan klaster lainnya. Dimana dapat diartikan bahwa pada klaster 4 terdapat 5 variabel yang memiliki nilai diatas rata-rata. Oleh karena itu, klaster 4 yaitu masuk ke dalam kategori daerah kriminalitas rawan karena memiliki nilai di atas rata-rata.

### 5. Perbandingan Model terbaik

Untuk mengetahui metode terbaik dari kedua metode tersebut maka dilihat dari nilai rasio simpangan baku dari nilai simpangan baku dalam klaster ( $S_w$ ) dan nilai simpangan baku antar klaster ( $S_B$ ).

Simpangan Baku	Ward	K-Medoids
Simpangan baku dalam klaster ( $\sigma_w$ )	337.79	332.24
Simpangan baku antar klaster ( $\sigma_B$ )	1222.24	1201.86
<b>Rasio simpangan baku klaster (<math>\sigma</math>)</b>	<b>0.276368</b>	<b>0.276434</b>

Berdasarkan table diatas, didapatkan nilai simpangan baku dalam klaster dan antar klaster untuk setiap metode. Memilih metode yang terbaik dengan melihat nilai simpangan baku dalam klaster

atau  $\sigma_w$  yang terkecil dan nilai simpangan baku antar klaster atau  $\sigma_B$  yang paling besar atau melihat nilai rasio simpangan baku yang terkecil. Berdasarkan tabel diatas dapat dilihat bahwa untuk metode *Ward* memiliki nilai rasio simpangan baku terkecil yaitu 0.276368. Oleh karena itu, metode *Ward* adalah metode yang terbaik dalam penelitian ini.

### KESIMPULAN

1. Jumlah kejadian kejahatan terhadap nyawa (X1) paling sedikit yaitu di provinsi DI Yogyakarta dan jumlah kejadian terbanyak di Provinsi Sumatera Selatan. Jumlah kejadian kejahatan terhadap fisik/badan (X2) paling sedikit yaitu di provinsi Kalimantan Utara dan jumlah kejadian terbanyak terbanyak di Provinsi Sumatera Utara. Jumlah kejadian kejahatan terhadap kesusilaan (X3) paling sedikit yaitu di provinsi Kalimantan Utara dan jumlah kejadian terbanyak di Provinsi Jawa Barat. Jumlah kejadian kejahatan terhadap kemerdekaan orang (X4) paling sedikit yaitu di provinsi Kalimantan Timur yaitu bernilai nol artinya tidak ada jumlah kejadian dan jumlah kejadian terbanyak di Provinsi Sulawesi Selatan. Jumlah kejadian kejahatan terhadap hak milik/barang dengan penggunaan kekerasan (X5) paling sedikit di provinsi Maluku Utara dan jumlah kejadian terbanyak di Provinsi Sumatera Utara. Jumlah kejadian kejahatan terhadap hak milik/barang tanpa penggunaan kekerasan (X6) paling sedikit di provinsi Maluku Utara dan jumlah kejadian terbanyak di Provinsi Sumatera Utara. Jumlah kejadian kejahatan terkait narkoba (X7) paling sedikit di provinsi Nusa Tenggara Timur dan jumlah kejadian terbanyak di Provinsi DKI Jakarta. Dan jumlah kejadian kejahatan terkait penipuan, penggelapan, dan korupsi (X8) paling sedikit di provinsi Maluku Utara dan jumlah kejadian terbanyak di Provinsi DKI Jakarta.
2. Hasil Pengklasteran dengan menggunakan metode *ward* dengan  $k=4$  yaitu : Klaster 1 yaitu daerah cukup aman memiliki anggota 10 Provinsi yaitu Provinsi Aceh, Jawa Tengah, Kalimantan Selatan, Kalimantan Timur, Lampung, Nusa Tenggara Timur, Papua, Riau, Sulawesi Tengah, dan Sulawesi Utara. klaster 2 yaitu daerah aman yang beranggotakan 17 Provinsi yaitu Provinsi Bali, Kepulauan Bangka Belitung, Banten, Bengkulu, Gorontalo, Papua Barat, Jambi, Kalimantan Barat, Kalimantan Tengah, Kalimantan Utara, Kepulauan Riau,

Maluku Utara, Maluku, Nusa Tenggara Barat, Sulawesi Barat, Sulawesi Tenggara, dan DI Yogyakarta. Kluster 3 yaitu daerah rawan yang beranggotakan 2 Provinsi yaitu Provinsi DKI Jakarta dan Sumatera Utara. Dan kluster 4 daerah cukup rawan yang beranggotakan 5 Provinsi yaitu Provinsi Jawa Barat, Jawa Timur, Sulawesi Selatan, Sumatera Barat, dan Sumatera Selatan. Hasil Pengklasteran dengan menggunakan metode *k-medoids* dengan menggunakan  $k=4$  yaitu : Kluster 1 yaitu daerah cukup aman memiliki anggota 8 Provinsi yaitu Provinsi Aceh, Jawa Tengah, Lampung, Nusa Tenggara Timur, Papua, Riau, Sulawesi Tengah, dan Sulawesi Utara. Kluster 2 yaitu daerah aman yang beranggotakan 19 Provinsi yaitu Provinsi Bali, Kepulauan Bangka Belitung, Banten, Bengkulu, Gorontalo, Papua Barat, Jambi, Kalimantan Barat, Kalimantan Selatan, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Kepulauan Riau, Maluku Utara, Maluku, Nusa Tenggara Barat, Sulawesi Barat, Sulawesi Tenggara, dan DI Yogyakarta. Kluster 3 yaitu daerah cukup rawan yang beranggotakan 2 Provinsi yaitu Provinsi DKI Jakarta dan Sumatera Utara. Dan kluster 4 yaitu daerah rawan yang beranggotakan 5 Provinsi yaitu Provinsi Jawa Barat, Jawa Timur, Sulawesi Selatan, Sumatera Barat, dan Sumatera Selatan.

### 3. Perbandingan metode terbaik

Nilai rasio simpangan baku yang terkecil terdapat pada metode *Ward* dengan nilai 0.276368. Sehingga metode terbaik yang dapat digunakan untuk melakukan pengklasteran dalam penelitian ini yaitu dengan menggunakan metode *Ward*.

### DAFTAR PUSTAKA

- Agusta, Y. (2007). K-Means - Penerapan, Permasalahan dan Metode Terkait Jurnal Sistem dan Informatika. *Vol 3*, 47-60.
- Arikunto, S. (2006). *Prosedur Penelitian : Suatu Pendekatan Praktik Edisi Revisi*. Jakarta: PT Rineka Cipta.
- Asia,S.E.(2018). *Crime Index by Country 2018*. Numbeo.  
[https://www.numbeo.com/crime/rankings\\_by\\_country.jsp?title=2020&region=035](https://www.numbeo.com/crime/rankings_by_country.jsp?title=2020&region=035)  
 (diakses pada tanggal 21 Juli 2020)
- Aziz, C. (2017). *Analisis Multivariat Clustering K-Means Untuk APBD Kecamatan-Kecamatan di Kota Semarang Tahun 2015*. Semarang: Universitas Negeri Semarang.
- Ediyanto, M. N. (2013). *pengklasifikasian karakteristik dengan metode k-means cluster analysis*. Buletin ilmiah Mat, Stat, dan Terapannya (Bimaster) , 1.
- Febriyana (2011). *Analisis Kluster K-Means dan K-Median Pada Data Indikator Kemiskinan*. Tangerang Selatan : Universitas Islam Negeri Syarif Hidayatullah Jakarta.
- Ghozali, I. (2011). *Aplikasi Analisis Multivariate dengan Program IBM SPSS 19*. Semarang: Undip.
- Gudono. (2014). *Analisis Data Multivariat*. Yogyakarta: BPFE.
- Han, J., & Kamber, M. (2006). *Data Mining : Concept and Techniques*. Waltham: Morgan Kauffman Publisher.
- <https://www.bps.go.id/publication/2019/12/12/>  
 (diakses pada tanggal 16 Juni 2020).
- Johnson, R.A. dan Wichern, D.W.. (2007). *Applied Multivariate Statistical Analysis 6 Edition*. Upper Saddle River, New Jersey 07458: Prentice Education, Inc.
- Laeli, S. (2014). *Analisis Cluster Average Linkage Method dan Ward's Method untuk Data Responden Nasabah Asuransi Jiwa Unit Link*. Yogyakarta : Universitas Negeri Yogyakarta.
- Mahmudi. (2013). *Clustering Mahasiswa Matematika Berdasarkan Karakteristik Menggunakan Metode K-Means*. Semarang: FMIPA Universitas Negeri Semarang.
- Nyoman Gde Prajnowiweka Ratmasa Taram, K. G. (2019). Pengklasteran Tingkat Kriminalitas Dengan Metode Agglomerative Dan K-Means Serta Peubah Pencirinya. *E-Jurnal Matematika Vol. 8(2)*, 102.
- Puspitasari, M.W dan Susanti, M. (2016). pengklasteran Kabupaten/Kota Berdasarkan Faktor-faktor yang mempengaruhi kemiskinan di Jawa Tengah Menggunakan Metode Ward dan Average Linkage. *Jurnal Matematika UNY*, 6.
- Santoso, S. (2015). *Menguasai Statistik Parametrik Konsep dan Aplikasi dengan SPSS*. Jakarta: PT Elex Media Komputindo.
- Santoso, S. (2014). *Statistik Multivariat Edisi Revisi konsep dan Aplikasi dengan SPSS*. Jakarta: PT Elex Media Komputindo.
- Setyawati, & Astri Widiastuti. (2017). *Implementasi Algoritma Partitioning Around Medoid (PAM) untuk Pengklasteran Sekolah Menengah Atas di DIY Berdasarkan Nilai Daya Serap Ujian Nasional*. Yogyakarta: Fakultas Sains dan Teknologi Universitas Sanata Dharma.

- Simamora, B. (2005). *Analisis Multivariat Pemasaran Edisi Pertama*. Jakarta: PT. Gramedia Pustaka Utama.
- Sirojuddin, A. (2016). *Analisis Custer pada Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Indeks Pembangunan Manusia*. Malang: Universitas Islam Negeri Maulana Malik Ibrahim.
- Statistik, B. P. (2019). *Statistik Kriminal 2019*. Jakarta: Badan Pusat Statistik.
- Sukestiyarno. (2013). *Olah Data Penelitian Berbantuan SPSS*. Semarang: Universitas Negeri Semarang.
- Supranto, J. (2004). *Analisis Multivariat Arti & Interpretasi*. Jakarta: PT RINEKA CIPTA.
- Vercillis, C. (2009). *Business Intelligence : Data Mining and Optimization for Decision Making*. Milan: WILEY.
- Widarjono, A. (2015). *Analisis Multivariat Terapan*. Yogyakarta: UPP STIM YKPN.

